COMPUTER VISION

Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works

An autonomous drone for search and rescue in forests using airborne optical sectioning

D. C. Schedl, I. Kurmi, O. Bimber*

Autonomous drones will play an essential role in human-machine teaming in future search and rescue (SAR) missions. We present a prototype that finds people fully autonomously in densely occluded forests. In the course of 17 field experiments conducted over various forest types and under different flying conditions, our drone found, in total, 38 of 42 hidden persons. For experiments with predefined flight paths, the average precision was 86%, and we found 30 of 34 cases. For adaptive sampling experiments (where potential findings are double-checked on the basis of initial classification confidences), all eight hidden persons were found, leading to an average precision of 100%, whereas classification confidence was increased on average by 15%. Thermal image processing, classification, and dynamic flight path adaptation are computed on-board in real time and while flying. We show that deep learning–based person classification is unaffected by sparse and error-prone sampling within straight flight path segments. This finding allows search missions to be substantially shortened and reduces the image complexity to 1/10th when compared with previous approaches. The goal of our adaptive online sampling technique is to find people as reliably and quickly as possible, which is essential in time-critical applications, such as SAR. Our drone enables SAR operations in remote areas without stable network coverage, because it transmits to the rescue team only classification results that indicate detections and can thus operate with intermittent minimal-bandwidth connections (e.g., by satellite). Once received, these results can be visually enhanced for interpretation on remote mobile devices.

INTRODUCTION

Use of uncrewed aerial vehicles (UAVs) by emergency services in civil applications is rapidly increasing (*1*), also for search and rescue (SAR) missions (*2*–*7*). In contrast to crewed aircraft, such as helicopters, drones are more flexible and cheaper in acquisition, maintenance, and operation; moreover, they avoid risks to pilots under difficult weather conditions. Considerable progress is being made to create fully autonomous drones that team up with action forces. However, technical and legal obstacles remain. For instance, under complex flying conditions aviation regulations require autonomous drones to support advanced safety features, such as traffic avoidance and flight termination systems, that are still under development for drones.

Effective imaging, for instance, is essential for autonomous UAVs. The narrow aperture optics of conventional cameras increase the depth of field and therefore project sharply the entire occlusion volumes (such as forests) into the images captured. Objects of interest at a particular distance often remain fully occluded. Wide-aperture optics (i.e., with diameters of several meters), which cause an extremely shallow depth of field, would be better suited to these cases but remain infeasible for airborne applications. Synthetic aperture (SA) sensing offers a means of overcoming these physical limitations; it is widely recognized to approximate theoretical wide-aperture sensors by computationally combining the signals of multiple small-aperture sensors or a single moving small-aperture sensor to improve resolution, depth of field, frame rate, contrast, and signal-to-noise ratio. This principle has been applied in a range of fields, for instance, for obtaining weather-independent images and reconstructing geospatial depth information using radar (*8*–*10*), for observing large celestial phenomena in outer space using radio telescopes (*11*, *12*), for reconstructing a defocus-free three-dimensional (3D) volume using interferometric microscopy (*13*), for generating high-resolution mapping of objects and seafloors using sonar (*14*–*17*), and for applying 2D SA imaging to shorter wavelengths (i.e., optical light) using light detection and ranging (LIDAR)/SA imaging laser (*18*, *19*). In the visible range, SA imaging (*20*–*27*) has been used to acquire structured light fields (regularly sampled multiscopic scene representations). Systems using large camera arrays capture 4D light ray data that are then used to support various digital postprocessing steps (such as refocusing, computation of virtual views with maximal SA, and varying depth of field) after their acquisition. In such systems, SA are constrained mainly by the physical size of the camera array used.

With airborne optical sectioning (AOS) (*28*–*34*), we have introduced a wide SA imaging technique that uses crewed or uncrewed aircraft, such as drones (Fig. 1A), to sample images within large (SA) areas from above occluded volumes, such as forests. On the basis of the poses of the aircraft during capturing, these images are computationally combined to integrate images. These integral images suppress strong occlusion and make visible targets that remain hidden in single recordings.

The computation of AOS integral images is based on unstructured light-field theory [refer to (*35*) and (*36*) for a thorough discussion], which represents image pixels as 4D light rays in a 3D volume. As illustrated in Fig. 1B, the origins of these rays are their corresponding cameras' position and their directions are determined by the aircrafts' poses and the cameras' intrinsic parameters.

Rays are lastly intersected with the surface of a digital elevation model (DEM; Fig. 1B). For rays that intersect at the same surface point, the associated pixel values are averaged. Repeating this for all rays and surface points results in an integral image. Thus, points on the ground appear in focus, whereas occluding structures (trees and vegetation) are defocused. In case a DEM is not available, the ground's surface is approximated by a focal plane whose parameters can be automatically determined (*33*).

Computer Science Department, Johannes Kepler University Linz, Linz, Austria.
*Corresponding author. Email: oliver.bimber@jku.at

Downloaded from https://www.science.org on October 19, 2022

Fig. 1. Autonomous drone supporting SAR missions. (**A**) Camera drone with the recording and processing equipment used in our experiments. In the case of a positive finding, an integral thermal image from the AOS approach, together with classification result and location are transmitted to a mobile device of the rescue team. (**B**) Basic principle of AOS. Movie S1 shows a video footage of the drone and sample test site.

In (*30*), we presented a statistical model to explain the efficiency of AOS with respect to occlusion density, occluder sizes, number of integrated samples, and size of the SA. Although the SA in (*30*) and (*34*) are 2D, the equations and AOS principles still hold in the case of 1D apertures as the ones being applied in this work. The main advantages of AOS over alternatives, such as LIDAR (*37*–*39*) or synthetic aperture radar (*8*–*10*), are its computational performance in real-time occlusion removal and its independence to wavelength. AOS has been applied in the visible spectrum (*28*) and in the far-infrared (thermal) spectrum (*31*) for wildlife observations (*32*) and search and rescue (SAR) (*34*). In addition, it can be applied to near-infrared wavelengths, for example, to address applications in agriculture or forestry. In (*34*), we demonstrated that integrating single images before classification rather than combining classification results of single images is substantially more effective when classifying partially occluded persons in aerial thermal images. Here, AOS was applied for image integration and occlusion removal. This finding, in principle, enables drone-supported SAR missions in heavily concealed areas, such as forests. In practice, however, several challenges remain to be overcome. Thus far, drones have been used only to sample images along a predefined flight path that covers a 2D SA, with all image processing done offline, after a flight, on high-performance computers. Both sampling of large 2D aperture areas and heavy image processing [in particular, computer vision–based pose estimation, which can take hours even on high-end graphics processing units (GPUs)] require a notable amount of time, which makes them impractical for time-critical applications, such as SAR. Furthermore, sampling along a fixed flight path does not allow regional differences in occlusion density that may require resampling to be considered.

Here, we report on two main contributions toward fully autonomous drones for SAR: First, we show that, compared with the classical approach of 2D aperture sampling and computationally expensive computer vision–based pose estimation, the effectiveness of person classification under occlusion conditions is not reduced when sampling along 2D SAs (i.e., 1D flight paths) and using imprecise Global Positioning System (GPS) and inertial measurement unit (IMU) measurements for pose estimation. Having resampled previous test flights (*34*), we report that imprecise 1D sampling of around 30 images and precise 2D sampling of around 300 images yield similar average precision (AP) scores of 92 to 93%. It follows that image processing can be implemented in its entirety using on-board mobile processors, thus enabling real-time classification during time-efficient flights. Second, we introduce an adaptive online sampling technique that changes the flight path dynamically during flight and uses classification confidences for decision-making. Thus, the drone can decide to resample a particular region if it receives even a weak hint from the classifier that a person might be hidden there. For path planning, we rely on a common potential field algorithm, as it supports iterative local planning at moderate computational costs. Our classification-driven resampling strategy, however, can be applied to other path planning techniques. We developed an autonomous prototype drone and evaluated it in the course of 17 field experiments over various forest types (conifer, broadleaf, and mixed) and under different flying conditions (daylight, temperature, and seasons).

With this prototype, we achieved an AP score of 86% for predefined flight paths and a 100% AP score for our adaptive sampling, where confidence scores improved by 15%, due to resampling. From a total of 42 hidden persons, 38 were found by the drone (30 of 34 with one false alert for experiments with predefined flight paths and 8 of 8 with no false alert for adaptive sampling experiments).

RESULTS

1D SA classification

In (*34*), we sampled 2D SA areas of 30 m by 30 m (this equals the ground coverage of our thermal camera's field of view at an altitude of about 35 m above ground level) with a sampling density of 1 m by 3 m. The high number of samples (about 300 images) and precise computer vision–based pose estimation ensured effective occlusion removal and a focused appearance of persons in the corresponding integral images (Fig. 2, A and B). A 1D SA sampling (i.e., integrating images captured along a 1D flight path) and image registration with instant but imprecise GPS and IMU measurements from the drone results in defocus (Fig. 2C). The defocused point spread of a person's thermal signal contains the forest's occlusion structure in its optical bokeh.

The advantage of 1D over 2D SA sampling is a substantial reduction in sampling time for covering the same area (by a factor of 10 in our example, because—at the same sampling rate—only 30 rather than 300 images are captured and integrated). The disadvantages are

Downloaded from https://www.science.org on October 19, 2022

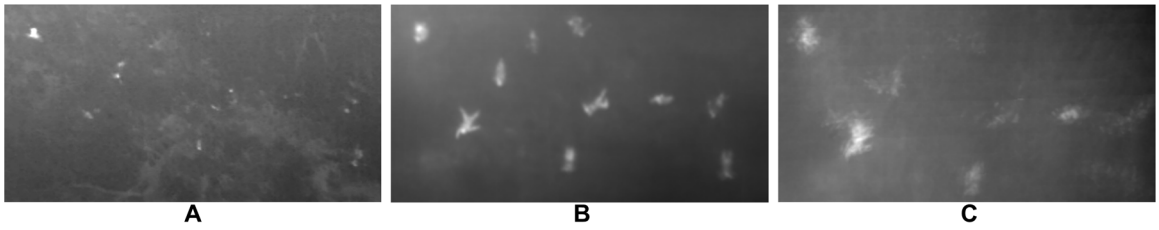

Fig. 2. 2D versus 1D SA imaging. The strong occlusion in single thermal images (**A**) can be removed effectively in integral images that densely sample 2D SA (**B**). Precise computer vision–based pose estimation leads to well-defined shapes of people. 1D SA sampling combined with GPS/IMU-based pose estimation leads to defocus in the resulting integral images (**C**).

less-efficient occlusion removal (due to undersampling) and defocus (due to less precise pose estimation). In particular, the latter has implications not only for classification itself but also for training a classifier. In (*34*), we showed that, in the case of efficient occlusion removal and precise image registration, the data needed for training a classifier are invariant to occlusion. Hence, training data can be recorded under control conditions in an open-field environment with the trained classifier still performing well for occluded test data recorded from above all types of forest. This is no longer the case for defocused 1D SA samplings, because the occlusion structure becomes an apparent part of each integral image's optical bokeh. Occlusion invariance can no longer be assumed, and, accordingly, training data must be produced in the presence of occlusion. However, the question arises of whether classification remains effective for cases such as that shown in Fig. 2C even if training data contain occlusion.

An initial experiment provided first evidence that this is the case: We split 11 test flights (F0 to F11, excluding F7) over various forest types (conifer, broadleaf, and mixed forest) from (*34*) into five training flights (F0, F4, F5, F10, and F11, with 132 labels of 19 persons), two validation flights (F1 and F8, with 65 labels of 10 persons), and four test flights (F2, F3, F6, and F9, with 204 labels of 26 persons). Details of these flights are provided in (*34*). Note that, because of its circular sampling, we did not use F7. The original 2D grid SA sampling of each flight was computationally resampled to sequences of 1D line SA samplings, whereas the measured GPS/IMU data were used for pose estimation in place of a computer vision–based pose estimation. Furthermore, a simpler classifier [you only look once (YOLOv4)–tiny (*40*)] that is suitable for less-performant mobile processors was used [previously, we applied the computationally more expensive YOLO3 classifier with spatial pyramid pooling (*41*, *42*)]. Results are summarized in Fig. 3 and presented in more detail in Table 1.

Figure 3 plots the overall (of all test flights combined) AP score over an increasing number (N) of integrated thermal images sampled along a 1D SA path. The AP metric (*43*) is commonly used to evaluate the performance of object classification algorithms; considers correct, wrong, and missed detections across multiple scenes; and is independent of the classifier sensitivity (i.e., does not require manual tuning of confidence thresholds). Samples were taken at intervals of 1 m. A maximum path length of 30 m (30 samples) was recorded in our flights. The results of discrete path length variations were fitted to a hyperbolic function (see the Average Precision Curve section of the Supplementary Materials) and extrapolated. This function indicates that high classification rates can be achieved for relatively small numbers of samples and short flight paths, even for 1D SA sampling with defocus and simpler classifiers. For example, an AP of 92.8% was achieved with $N = 30$. In comparison, for the 2D SA sampling with better focus and classifier, an AP of 92.2% was achieved with $N > 300$ (*34*). Note that we augmented the training data of our new classifier with all path length variations (from $N = 1$ to 30). This finding suggests that neither 2D SA sampling nor precise pose estimation is necessary for achieving similar classification performance at a substantially higher speed.

Training data, however, are no longer invariant to occlusion. When trained with occlusion data, an AP of 62.2% is achieved for single-image SAs ($N = 1$; pinhole). Note that the reported single-image AP score, when trained on nonocclusion training data, is 24.8% in (*34*). This indicates that single-image classification improves when trained with occlusion data, as the classifiers might learn in addition from occlusion patterns. The low detection scores and the high number of false positives (FPs; see Table 1), however, still make single-image classification impractical for real-world applications.

Figure 3 also illustrates that the steep increase in classification performance followed by a flattening beyond a small N is in line with the increase in visibility improvement in integral images for increasing N. This was confirmed by our statistical model described in (*30*), where we claimed that the performance of SA imaging for occlusion removal does not increase infinitely with increasing apertures and higher sampling rates but that there is a (relatively low) limit in both for achieving a maximum improvement. These low limits make AOS practical for time-critical applications, such as SAR.

Predefined search

In the second series of experiments, we validated our findings in practice in the course of eight new test flights (covering a total area of about 6 ha) over various forest types, where all processing was carried out in real time on the drone and during flight but where the flight paths were still predefined and static (following fixed waypoints). Here, we used test flights from (*34*) for training (F0, F2 to F6, F8, F10, and F11) and training validation (F1 and F9) and, again, augmented with path length variations from $N = 1$ to 30. During each new test flight, the drone sampled segments of 30-m-long 1D SAs with 1 sample/m along defined waypoints. Thus, after each segment, an integral image was computed and classified. Table 2 presents the results, whereas Fig. 4 illustrates them for the example of flight F16. Of 34 persons present, 30 persons were found, and only one person was incorrectly detected. The overall AP score was lower than in the first experiments (86% versus 93%). This is due to Table 1 comparing test flights of our previous offline classification study presented in (*34*), where compass errors were corrected manually after recording. The data of the new test flights in Table 2 were computed fully automatically during flights and contain compass

Downloaded from https://www.science.org on October 19, 2022

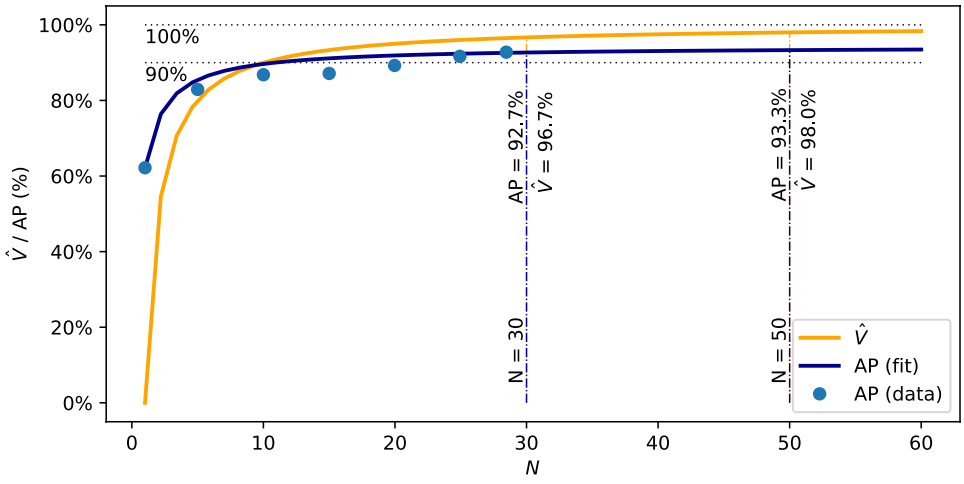

Fig. 3. 1D SA classification performance. Classification performance increases for all test flights in Table 1 (blue) and modeled visibility improves (*30*) (yellow) with increasing number of integrated sample images *N*. Movie S2 illustrates the behavior of the blue curve for integral images of four sample flights.

errors. If compass errors are manually corrected after the flights, then the AP score of Table 2 increases to 92% (shown in parentheses). Thus, better compass modules will improve the classification results.

Many real-time object detection algorithms that are based on convolutional neural networks, such as YOLO (*44*), divide images into regions, for each of which they predict bounding boxes and probabilities (confidence scores). These confidence scores are thresholded for making a final classification decision. Depending on the choice of threshold, either too many FPs (incorrect detections) or too few true positives (TPs; i.e., persons are missed) can be the result. For both series of experiments (Tables 1 and 2), we selected a YOLO confidence score threshold of 10%. This minimizes FPs. A lower confidence score would increase the number of incorrectly detected persons (false alerts) but might also increase correct detections.

Because confidence scores correlate directly with the amount of occlusion, areas with detections that indicate the presence of a person should be resampled (double-checked) to observe the development of the confidence score before making a final decision. If an initially low score drops further after resampling, then this confirms an FP detection. If an initially high score increases, then this confirms a TP detection. However, this is only possible if the flight path is adapted dynamically on the basis of current classification results.

Classification-driven adaptive search

In the third series of experiments, we dynamically computed and adapted the flight path on board the drone rather than have it follow a predefined set of fixed waypoints. Here, the goal was to find a person as quickly and as reliably as possible, allowing the drone to make the decisions autonomously on how and where to search. Because the drone's flight path changes dynamically based on classification results, we chose to build on common potential field–based path planning (*45*–*50*). Because of its scalability, limited computational complexity, and adaptability, it has been used previously in the course of real-time adaptive path planning for obstacle avoidance (*48*) and target detection and tracking (*49*, *50*). The interested reader is referred to summaries on path planning for UAVs (*46*, *47*, *51*–*61*).

The search areas were split into grids of 30 m–by–30 m cells, and each cell was initialized with the rescue team's estimate of the likelihood of finding a person within it. Each cell was scanned centered (horizontally or vertically) from edge to edge with a 30-m SA, as in all previous experiments. The decision on which cells are scanned and in which order was based on the cell probabilities and by evaluating the following potential field equation [as explained in (*45*)] for each cell

$$f(i) = P(i) * e^{-\|x - c_i\|_2} \quad (1)$$

where $P(i)$ is the ith cell's probability, x is the current location of the drone, and c_i is the center position of cell i. The cell scanned next was that with the highest potential $f(i)$. If multiple cells had the same maximum $f(i)$, then we selected (among them) the cell that had the highest potential density in its neighborhood (successively enlarging the neighborhood until a maximum potential density cell was identified)

$$P_d(i) = \sum_{j \neq i} \frac{P(j)}{\|c_i - c_j\|_2} \quad (2)$$

where j are the neighbors of i.

The probabilities of scanned cells were set to zero to avoid revisiting them. We stopped scanning under three conditions: (i) if a confirmed TP was detected, (ii) if a maximum path length (maximum flying time) was reached, and (iii) if the entire search region had been covered. Figure 5A illustrates an example path that was computed dynamically for a search area with an initial probability map. By evaluating Eqs. 1 and 2, we maximized the likelihood of detection while minimizing search time based on the rescue team's initial assessment of the situation.

If a detection was made (we considered a detection to be weak if it had a confidence score of at least 5%), then resampling the corresponding cell was triggered. Because the orientation of occluders is unknown, we resampled with an SA that was orthogonal to the previous one to increase the likelihood of visibility (in case of nonuniform occlusion distributions). Because the sampling rate of integral images is not uniform (lower in the periphery than in the center), this new SA was not cell centered but centered over the detection to achieve the highest possible sampling rate as close as possible to the target. Note that although combining the two samplings (cell centered and detection centered) seems to be an option, this does not lead to an improvement of the classification result because, as shown earlier, classification is unaffected by the SA's sampling dimensionality. Therefore, the classifier is trained with 1D SA sampling data. Classifying 2D sampled data with it leads to weaker results. The reason is that, although the appearance of focused targets in integral images is similar in both cases, the bokeh of defocused occluders depends on the sampling pattern. The latter will affect the classification results. Thus, the sampling dimensionality of the training data and the test data must match.

We replace any cell-centered scores by the detection-centered scores. In all nine experiments of Table 3, this strategy leads to an

Downloaded from https://www.science.org on October 19, 2022

Table 1. 1D SA classification performance. Classification results of test flights from (*34*), resampled to 1D SA segments, for various path lengths *N*. GT is the number of ground truth labels, TP are the true positives, FP are the false positives, and AP is the AP score. F9 is a flight over an empty (no hidden persons) forest. Note that, because of resampling, the path lengths per flight may vary slightly. Thus, *N* represents the average. Note also that fewer segments could be resampled for F2 (*N* = 28.46) than for the other flights. This also led to fewer GT labels (70 instead of 86) in this case, thus indicated with asterisks above TP and FP. Movie S1 visualizes the classification improvement for increasing *N*. Figure S1 provides satellite and color images of the corresponding test sites. n/a, not applicable.

	N = 1 (pinhole)			*N* = 5			*N* = 10			*N* = 15			*N* = 19.97			*N* = 24.93			*N* = 28.46		
Flight (GT)	**AP**	**TP**	**FP**	**AP**	**TP**	**FP**	**AP**	**TP**	**FP**	**AP**	**TP**	**FP**	**AP**	**TP**	**FP**	**AP**	**TP**	**FP**	**AP**	**TP***	**FP***
F2 (70–86)	64%	65	80	90%	78	16	94%	81	8	96%	80	12	97%	81	7	96%	81	5	97%	66	5
F3 (53)	36%	24	42	46%	30	31	52%	31	24	55%	32	21	60%	37	19	72%	40	19	74%	43	28
F6 (65)	86%	56	16	99%	64	1	99%	64	4	98%	64	2	98%	64	1	98%	64	1	98%	64	2
F9 (0)	n/a	0	6	n/a	0	7	n/a	0	2	n/a	0	4	n/a	0	2	n/a	0	2	n/a	0	2
All (188–204)	62%	145	144	83%	172	55	87%	176	38	87%	176	39	89%	182	29	92%	185	27	93%	173	37

increase in confidence for correct person detections [persons found (PF)] and to a reduction in confidence for incorrect detections [persons incorrectly found (PI)]. For multiple detections within the same cell, we scanned multiple SAs accordingly. Figure 5B presents an example showing the subarea highlighted in Fig. 5A in which two detections were made at one border of the cell (with confidence scores of 11 and 27%). After resampling, an increase in the confidence score of one detection (from 27 to 51%) confirmed a TP, whereas a drop in the confidence score of the other detection (from 11 to 0%) confirmed a FP. Figure 5C illustrates another example, where a weak 7% confidence detection was resampled and, based on a 0% confidence score, confirmed to be a FP.

Table 3 summarizes the results of nine adaptive search test flights, where eight of eight persons have been found and because of adaptive sampling, no incorrect persons were detected. On average, resampling increased the confidence scores by around 15% for confirmed TPs (PF) and decreased confidence scores by around 16% for confirmed FPs (PI). With a 10% confidence threshold as used before, and without adaptive resampling, false decisions would have been made in four of the nine flights: Two persons would not have been detected, and two erroneous detections would have occurred.

DISCUSSION

For automatic person classification in occluding forests, 1D SA imaging performs, despite GPS errors, equally well as 2D SA imaging of 10 times the number of samples combined with time-consuming precise pose estimation. Nevertheless, image integration remains necessary to achieve effective classification rates (62% AP for single images and 93% AP for integrals with 30 images; see Table 1). This finding motivated the development of an autonomous drone for SAR missions in forests, which adapts its flight path to achieve real-time classification results. In the course of 17 field experiments over various forest types (conifer, broadleaf, and mixed), we show that our prototype finds humans reliably. From a total of 42 hidden persons, 38 were found (30 of 34 for predefined flight paths and 8 of 8 for adaptive sampling experiments).

Unlike video streaming and remote processing, on-board processing does not require fast and stable network coverage, which might not always be available for SAR missions in remote areas. Intermittent minimal-bandwidth satellite connections are sufficient to transmit classification results of detections only. In our case, a search mission requires only an (open access) DEM and an (optional) initial probability map. Both can be uploaded on site before takeoff.

False-negative classifications (person not found) are equally problematic as FPs (person wrongly found) because the latter might lead to an unnecessary moving out of the rescue team to the wrong place. Therefore, we allow the rescue team to make the final decision based on the transmitted integral images of detections to a mobile ground station (Fig. 1A). Because of GPS pose error–related misregistration, however, these images are not easy to interpret visually. Automatic pose error reduction can be applied in the course of postprocessing to achieve substantial improvements, as explained in (*62*). These improvements, however, will not be beneficial to the automatic person classification itself because small regions of interest around the detections are required for visual optimization. The regions of interest, however, are the result of the classification. Registration optimization for entire images is not feasible because processing time and error proneness would be too high. Thus, visual enhancement of received classification results can be computed on remote mobile devices to support the rescue team in decision-making. Figure S4 illustrates the results of offline registration enhancement for all correct person detections (PF) in Table 3.

Carrying out experiments to gain practical experience together with emergency forces (such as firefighters, mountain rescue services, police, or armed forces) is on our agenda for future work. On the basis of the outcome of these experiments, the sensitivity (i.e., the confidence threshold) of the classifier can then be fine-tuned on demand.

Thermal imaging has limitations when applied in warm surrounding environments that affect the performance of AOS. To ensure a measurable temperature difference between bodies and environment in thermal images, flights F0 to F28 were performed from October to April. Only direct sunlight led to heat spots on trees and on the ground. Person classification in single thermal images, however, has been shown to be working in warm environments [e.g., for pedestrian detection (*63*)] if a temperature difference is resolvable by the camera. In this case, AOS will also be successful. However, SAR missions are more critical in cold conditions because missing people might freeze to death if not found in time.

Because of current aviation regulations, all test and training flights required visual line of sight. Short battery life restricted flying times to a maximum of 15 to 20 min. Professional drones with

Downloaded from https://www.science.org on October 19, 2022

Table 2. Predefined search experiments. Results of test flights with predefined flight paths, including GPS coordinates, date of flight, path length, and forest type at test sites. AP, the AP score; PP, the number of persons present at the test site; PF, the number of persons correctly found; and PI, the number of incorrect person detections (false alerts). Note that flight paths were divided into 30-m segments (and residues). Note also that, because of overlapping fields of view of the integral images, the same person may be labeled multiple times. Because it matters whether—and not how many times—a person was detected during the flight, we report person-specific PP, PF, and PI figures rather than label-specific GT, TP, and FP values as in Table 1. The numbers in parenthesis indicate scores after manual (offline) compass correction, leading to improved AP scores. Note, however, that a manual compass correction cannot be applied during flights. The results of F16 are shown in Fig. 4**C**, and the results of all other flights are illustrated in fig. S2. Aerial RGB samples of the corresponding test sites are shown in fig. S1.

ID	Latitude	Longitude	Date	Length	Forest	AP	PP	PF	PI
F12	48.33279428	14.33015898	13 Oct 2020	119.4 m	Conifer	96.7% (100.0%)	2 (2)	2 (2)	1 (0)
F13	48.33295299	14.33105885	13 Oct 2020	263.9 m	Broadleaf	100.0% (100.0%)	2 (2)	2 (2)	0 (0)
F14	48.33292935	14.33053121	15 Oct 2020	388.0 m	Conifer, broadleaf	100.0% (100.0%)	6 (6)	6 (6)	0 (0)
F15	48.33292935	14.33053121	20 Oct 2020	394.8 m	Conifer, broadleaf	75.6% (76.9%)	6 (6)	5 (5)	0 (0)
F16	48.33980546	14.33186291	22 Oct 2020	182.3 m	Mixed	91.1% (100.0%)	6 (6)	6 (6)	0 (0)
F17	48.33980546	14.33186291	22 Oct 2020	180.0 m	Mixed	63.3% (83.3%)	6 (6)	4 (4)	0 (0)
F18	48.3400508	14.33236699	28 Oct 2020	181.8 m	Mixed	100.0% (100.0%)	2 (2)	2 (2)	0 (0)
F19	48.3400508	14.33236699	28 Oct 2020	211.9 m	Mixed	75.0% (95.0%)	4 (4)	3 (3)	0 (0)
All				1922.1 m		86.2% (92.2%)	34 (34)	30 (30)	1 (0)

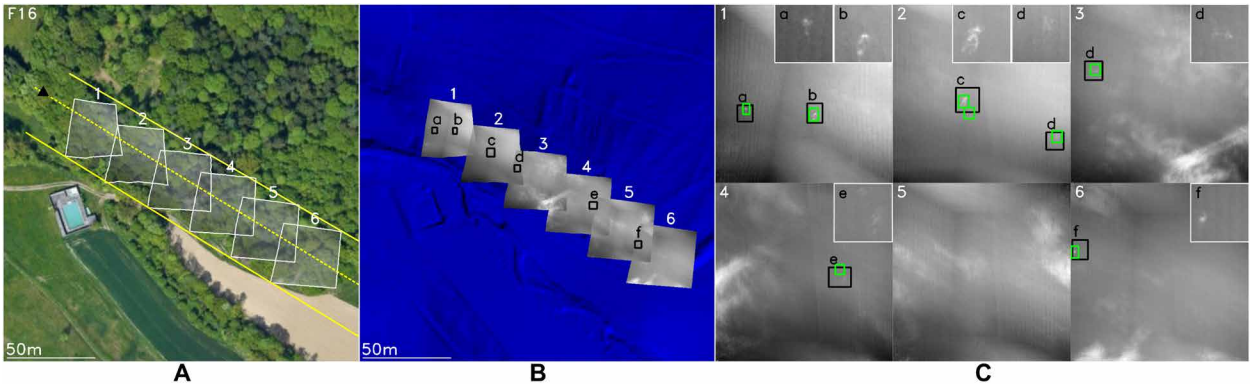

Fig. 4. Predefined search. Visualization of the F16 test flight from Table 2 (see also movie S3). The flight path was predefined [dotted yellow line in (A)]. The field of view [solid yellow lines in (A)] along the path was split into 30-m segments (and residue). An integral image was computed for each segment [thermal image overlays in (A) and (B)]. (**A**) A satellite view and (**B**) the DEM of the test site. (**C**) Integral images and close-ups with locations at which persons are present (black boxes) and classifications results where persons are detected (green boxes). Note that the drone's orientation was kept constant (facing north) to reduce dynamic compass errors during flight.

combustion engines (e.g., boxer or wankel) are more suitable for SAR missions because they can fly for up to 6 hours and carry payloads of up to 75 kg. Now, we are investigating a new long-endurance prototype together with drone manufacturers and emergency forces, which will also apply a high-quality real-time kinematic GPS system (see fig. S5). Appropriate legal regulations for autonomous flights beyond visual line of sight are in development and will most likely be approved much earlier for emergency operations (SAR, firefighting, and disaster management) than for other applications.

For autonomous flights beyond visual line of sight, additional components are essential, such as collision detection and avoidance (ground obstacles, such as power lines or other aircrafts). In areas where no reliable DEM is available, advanced ground obstacle avoidance will be necessary, and integral images have to be computed on the basis of automatically adjusted focal planes (*33*). Furthermore, visual search techniques, such as AOS, are complementary to mobile phone tracking. Although the latter is good for a coarse localization of a radio signal, AOS supports exact positioning and

Downloaded from https://www.science.org on October 19, 2022

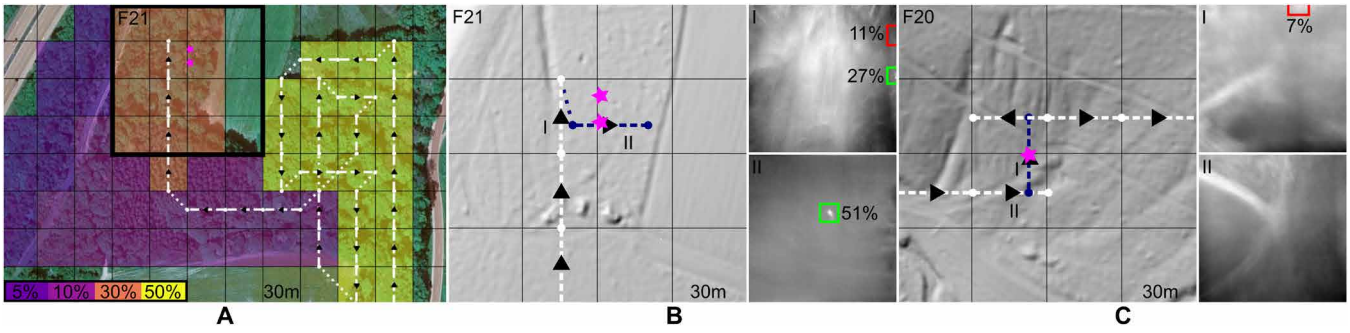

Fig. 5. Adaptive search. Satellite image of the test site overlaid with the flight path computed dynamically based on potential fields (**A**). The initially provided probabilities are color coded. Dashed white lines indicate scanned SAs within 30 m–by–30 m cells (1 m/s). Dotted white lines indicate fast flight segments (3 m/s) without sampling. Black triangles, the orientations of which illustrates the flight direction, mark positions of computed integral images. DEM overlaid with the close-up subarea of flight F21 (**B**). The two detections (pink stars) made in the initial integral image I are confirmed to be one correct (green box) and one incorrect (red box) classification in the resampled integral image II. Resampled SAs and fast flight segments without sampling are indicated by dashed and dotted black lines, respectively. The case of an empty scene in which an incorrect detection was confirmed (F20 in Table 3) is shown in (**C**). See also movie S4.

Table 3. Adaptive search experiments. Results of test flights with adaptive flight paths, including GPS coordinates, date of flight, path length, and forest type at test sites. ΔC is the change in confidence score from the initial sampling of a cell to resampling of the corresponding detections for PF and PI. (C) is the final confidence score after resampling. Cases highlighted in bold would have led to misclassifications (persons not found or false alerts) for a confidence threshold of 10%. PP, PF, and PI are after resampling. Figure 5 (B and C) shows cases F20 and F21, and fig. S3 illustrates all other cases. Figure S1 provides satellite and aerial RGB samples of the corresponding test sites.

ID	Latitude	Longitude	Date	Forest type	PF ΔC (C)	PI ΔC (C)	PP	PF	PI
F20	48.3398254	14.3327864	27 Nov 2020	Mixed	Empty	−6.7% (0.0%)	0	0	0
F21	48.3338445	14.3301653	30 Nov 2020	Mixed	23.9% (51.1%)	**−10.4% (0.4%)**	1	1	0
F22	48.3338445	14.3301653	30 Nov 2020	Mixed	8.4% (44.2%)	None	1	1	0
F23	48.3327007	14.330069	1 Dec 2020	Conifer	15.5% (38.7%)	**−40.4% (0.1%)**	1	1	0
F24	48.3332141	14.3314393	1 Dec 2020	Broadleaf	**16.7% (22.6%)**	None	1	1	0
F25	48.3332141	14.3314393	2 Dec 2020	Broadleaf	2.6% (35.6%)	None	1	1	0
F26	48.3327007	14.330069	3 Dec 2020	Conifer	**19.6% (19.6%)**	−6.9% (0.0%)	1	1	0
F27	48.3327007	14.330069	4 Dec 2020	Conifer	18.2% (39.5%)	None	1	1	0
F28	48.3338445	14.3301653	4 Dec 2020	Mixed	11.3% (43.6%)	None	1	1	0
All					14.5%	−16.1%	8	8	0

identification. Optimizing the data structure of our DEM to support efficient scalability through tiling is on our agenda for future work.

The thermal camera we used supports capture rates of 1 image/s. For 1 sample/m, this restricted flying speed to 1 m/s and integral image computation and classification to discrete 30-m steps (fast flight segments without sampling were done with 3 m/s). Faster imaging solutions would enable higher flying speeds (i.e., larger search areas) and continuous integral image updating and classification after capturing each image (e.g., after every meter flown). Early experiments indicate already that we can achieve a 4 to 10 m/s of sampling with our prototype by using a fast video-grabbing hardware instead of a direct image download from the camera.

Potential field–based path planning makes local decisions and does not lead to globally optimal solutions. However, global path planning algorithms (*52*)—such as evolutionary algorithms, rapidly exploring random trees, Voronoi diagrams, A* algorithms, and others—are computationally too complex and will not run in real time on a mobile processor for realistic scales. A good trade-off between real-time dynamic adaptability and considering larger neighborhoods for path planning must be investigated. Upgrading to a faster mobile processor or vision processor (e.g., NVIDIA Jetson or Google Coral) might support more advanced path planning algorithms. Adaptive search algorithms based on lost person movement modeling (*64*–*67*) needs to be considered as well in the future. Furthermore, multidrone solutions could reduce search time even further. All of these improvements, including enhancement of occlusion removal and classification and application to other fields (such as wildlife observation, finding fire hot spots, surveillance, and border control), fall within the scope of our future work. We believe that autonomous drones such as ours will play an essential role in human-machine teaming of rescue missions and emergency operations to come.

MATERIALS AND METHODS

We equipped an octocopter (MikroKopter Okto XL 6S12; 945-mm diameter, about 4.5 kg, and two lithium ion polymer (LiPo)

Downloaded from https://www.science.org on October 19, 2022

4500-mAh batteries) with a thermal camera (FLIR Vue Pro; 9-mm fixed focal length lens, 7.5- to 13.5-μm spectral band, 14-bit nonradiometric, and 118 g), a single-board system-on-chip computer (SoCC) (RaspberryPi 4B; 5.6 cm by 8.6 cm, 65 g, and 8-gigabyte (GB) random-access memory), an LTE communication hat (Sixfab 3G/4G and long-term evolution (LTE) base hat and a subscriber identification module (SIM) card; 5.7 cm by 6.5 cm and 35 g), and a vision processing unit (VPU) (Intel Neural Compute Stick 2; 7.2 cm by 2.7 cm by 1.4 cm and 30 g). The equipment (total weight of 320 g) was mounted on a rotatable gimbal, and the camera was pointed downwards during flight, as shown in Fig. 1.

The SoCC established communication with the drone (receiving IMU/GPS positions and sending waypoint instructions including GPS location, orientation, and speed) via a serial protocol, triggered the thermal camera using a pulse width–modulated signal and its purpose input-output pins (general purpose input/output), downloaded the recorded images from the camera's memory, preprocessed the images, and computed the integral image during flight. Our prototype is equipped with a nondifferential GPS with a 2.5-m horizontal positional accuracy at 50% circular error probable. Preprocessing involved adjusting the mean of the thermal images, removing the lens distortion using OpenCV's pinhole camera model, and cropping the images to a field of view of 50.82° and a resolution of 512 pixels by 512 pixels. Preprocessing and integration of 30 thermal images required around 748 and 90 ms on the SoCC. The VPU was initialized with our pretrained network weights for object detection using YOLOv4-tiny (*40*) at start-up and detected persons in the integral images while flying. Detections were computed in 84 ms per integral image and were transmitted via an LTE connection to a remote mobile device using the communication hat. The main software running on the SoCC was implemented using Python, whereas submodules, such as the drone communication protocol and integral computation, were implemented using C, OpenGL, and C++ and integrated using Cython. How integral images are computed was explained in detail in (*28*, *30*, *34*). On the SoCC, integrals are computed by projecting and averaging all single images on the ground surface (i.e., DEM or focal plane) using OpenGL's projective texture mapping technique (*68*). This work differs from (*28*, *30*, *34*) in that we did not assume the ground surface to be planar but approximated it with a DEM. The DEMs for the test sites are available freely from the Upper Austrian State Government (*69*) in a common GeoTIFF format. We converted a predefined region to a triangle mesh using the geospatial data abstraction library (GDAL) software (*70*), requiring 4 s for a 100 m by 100 m of DEM and 10 s for a 300 m by 300 m of DEM on a standard PC. For the experiments in Tables 2 and 3, the DEM was used to calculate the flying altitude of the drone. Thus, static obstacles and height variations in the DEM (e.g., hills or valleys) are considered. However, this does not support collision detection and avoidance for obstacles not present in the DEM (e.g., power lines or other aircrafts). For this, commercial collision detection systems [such as FLARM (*71*) or automatic dependent surveillance-broadcast (ADS-B) (*72*)] can be integrated. The central position of each SA was used to compute the integral images, within which persons were labeled if visible. The resolution of all integral images was always 512 pixels by 512 pixels.

For training and validation datasets, the following augmentations were applied. We randomly rotated the integral images (original and nine random rotations), changed the altitude of the DEM from −3 to +3 m in steps of 1 m (seven variations), and varied the number of image samples per SA (pinhole, 1 to 5, 5 to 10, 10 to 15, 15 to 20, 20 to 25, and 25 to 30), thus obtaining a total of 490 augmentations per SA. The field of view per SA and the labels of hidden persons were kept the same when varying the number of images. This means, for example, that only one image per SA was used for $N = 1$ (see movie S2). Note that 6 of 115 individual SAs were shorter than 25 cameras, and 1 SA was shorter than 20 cameras, which resulted in fewer long SAs.

For training our object detector with the YOLOv4-tiny (*40*) network architecture, we classified a single object class (person) in the integral images. The initial 29 network layers (of 39 layers) were pretrained with the ImageNet (*73*) dataset. We used the Darknet software for training and set the starting learning rate and the batch and subdivision sizes to 0.00261, 64, and 16, respectively. Training weights were stored after every 200 iterations and evaluated on the training set. The weights with the highest AP score [intersection over union (IoU) of 25%] were used in the subsequent experiments (Tables 1 to 3). The FP, TP, and AP scores in Table 1 were computed in MATLAB with an IoU threshold of 25%. For the FP and TP scores, detections below 10% confidence were discarded.

For our experiments summarized in Table 1, we reused the dataset from (*34*) and computationally resampled the recordings to sequences of 1D line SAs using the original GPS/IMU measurements. To make the new results comparable with the old 2D SA results in (*34*), we retained the manual compass correction as described in (*34*). For all other experiments, a manual compass correction was not possible, because images and drone poses were processed immediately in real time during flight. For the experiments in Table 1, we split 11 test flights (F0 to F11, excluding F7) from (*34*) into five training flights (F0, F4, F5, F10, and F11), two validation flights (F1 and F8), and four test flights (F2, F3, F6, and F9). The classifier for the additional test flights F12 to F28 (in Tables 2 and 3) was trained using flights F0 to F11 (excluding F1, F7, and F8) for training and F1 and F8 for validation.

For the experiments summarized in Table 2, the drone followed a preplanned path, recorded images, computed integral images, and performed detections after every 30-m SA segment during flight. After a flight, persons were labeled, and AP, PF, and PI scores were computed. For accurate labeling, we asked individuals to record their GPS locations on site using their smartphones. Because of the overlapping fields of view, the same person might appear in multiple integral images. This is considered in the PF and PI metric in Table 2. For PF, we considered a person as found if at least one detected a bounding box (with at least 10% confidence score) overlapped with the person's ground truth (GT) label. Because of the GPS/IMU pose errors, we applied an IoU threshold of 1% and ignored multiple detections in the same GT bounding box. We classified a detection as incorrect (PI) if it had a confidence score of more than 10% but was outside of all GT bounding boxes. Furthermore, to compare scores of Tables 1 and 2, we additionally applied manual compass corrections offline to update results of Table 2 (simulating better compass measurements). The compass corrections lead to an improvement of confidence scores and, hence, to an increase in AP. This indicates that more precise compass modules improve the classification during flight.

For the experiments summarized in Table 3, the potential field equations (Eqs. 1 and 2) were evaluated on the SoCC to decide on the next target cells based on the drone's current position, the probability map (predefined and updated after visiting each cell), and

Downloaded from https://www.science.org on October 19, 2022

previous detection result. Computation time for planning the next target cell was 1 ms.

SUPPLEMENTARY MATERIALS

robotics.sciencemag.org/cgi/content/full/6/55/eabg1188/DC1
Section S1
Figs. S1 to S5
Movies S1 to S4

REFERENCES AND NOTES

1. H. Shakhatreh, A. H. Sawalmeh, A. Al-Fuqaha, Z. Dou, E. Almaita, I. Khalil, N. S. Othman, A. Khreishah, M. Guizani, Unmanned aerial vehicles (UAVs): A survey on civil applications and key research challenges. *IEEE Access.* **7**, 48572–48634 (2019).
2. W. Sonia, N. Trigoni, Supporting search and rescue operations with UAVs, in *Proceedings of the 2010 International Conference on Emerging Security Technologies*, 6 to 7 September 2010, Canterbury, UK, pp. 142–147.
3. M. A. Goodrich, B. S. Morse, D. Gerhardt, J. L. Cooper, M. Quigley, J. A. Adams, C. Humphrey, Supporting wilderness search and rescue using a camera-equipped mini UAV. *J. Field Robot.* **25**, 89–110 (2008).
4. T. Teodor, K. Schmid, P. Lutz, A. Domel, M. Kassecker, E. Mair, I. L. Grixa, F. Ruess, M. Suppa, D. Burschka, Toward a fully autonomous UAV: Research platform for indoor and outdoor urban search and rescue. *IEEE Robot. Autom. Mag.* **19**, 46–56 (2012).
5. M. Silvagni, A. Tonoli, E. Zenerino, M. Chiaberge, Multipurpose UAV for search and rescue operations in mountain avalanche events. *Geomatics Nat. Hazards Risk* **8**, 18–33 (2017).
6. V. A. Feraru, R. E. Andersen, E. Boukas, Towards an autonomous UAV-based system to assist search and rescue operations in man overboard incidents, in *Proceedings of the 2020 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, Abu Dhabi, United Arab Emirates, 4 to 6 November 2020, pp. 57–64.
7. A. Birk, B. Wiggerich, H. Bülow, M. Pfingsthorn, S. Schwertfeger, Safety, security, and rescue missions with an unmanned aerial vehicle (UAV). *J. Intell. Robot. Syst.* **64**, 57–76 (2011).
8. A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, K. P. Papathanassiou, A tutorial on synthetic aperture radar. *IEEE Geosci. Remote Sens. Mag.* **1**, 6–43 (2013).
9. C. J. Li, H. Ling, Synthetic aperture radar imaging using a small consumer drone, in *Proceedings in 2015 IEEE International Symposium on Antennas and Propagation &USNC/URSI National Radio Science Meeting*, Vancouver, BC, Canada, 19 to 24 July 2015, pp. 685–686.
10. P. A. Rosen, S. Hensley, I. R. Joughin, F. K. Li, S. N. Madsen, E. Rodriguez, R. M. Goldstein, Synthetic aperture radar interferometry. *Proc. IEEE* **88**, 333–382 (2000).
11. R. Levanda, A. Leshem, Synthetic aperture radio telescopes. *IEEE Signal Process. Mag.* **27**, 14–29 (2010).
12. D. Dravins, T. Lagadec, P. D. Nuñez, Optical aperture synthesis with electronically connected telescopes. *Nat. Commun.* **6**, 6852 (2015).
13. T. S. Ralston, D. L. Marks, P. S. Carney, S. A. Boppart, Interferometric synthetic aperture microscopy. *Nat. Phys.* **3**, 129–134 (2007).
14. R. E. Hansen, in *Introduction to Synthetic Aperture Sonar in Sonar Systems* N. Kolev, Ed. (IntechOpen, 2011), pp. 3–28.
15. M. P. Hayes, P. T. Gough, Synthetic aperture sonar: A review of current status. *IEEE J. Ocean. Eng.* **34**, 207–224 (2009).
16. R. E. Hansen, H. J. Callow, T. O. Sabo, S. A. V. Synnes, Challenges in seafloor imaging and mapping with synthetic aperture sonar. *IEEE Trans. Geosci. Remote Sens.* **49**, 3677–3687 (2011).
17. H. Bülow, A. Birk, Synthetic aperture sonar (SAS) without navigation: Scan registration as basis for near field synthetic imaging in 2D. *Sensors* **20**, 4440 (2020).
18. Z. W. Barber, J. R. Dahl, Synthetic aperture ladar imaging demonstrations and information at very low return levels. *Appl. Opt.* **53**, 5531–5537 (2014).
19. S. Turbide, L. Marchese, M. Terroux, A. Bergeron, Synthetic aperture lidar as a future tool for earth observation. *Proc. SPIE* **10563**, 105633V (2017).
20. V. Vaish, B. Wilburn, N. Joshi, M. Levoy, Using plane + parallax for calibrating dense camera arrays, in *Proceedings of the 2004 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Washington, DC, USA, 27 June to 2 July 2004, pp. 2–9.
21. V. Vaish, M. Levoy, R. Szeliski, C. L. Zitnick, S. B. Kang, Reconstructing occluded surfaces using synthetic apertures: Stereo, focus and robust measures, in *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, New York, NY, USA, 22 June 2006, pp. 2331–2338.
22. H. Zhang, X. Jin, Q. Dai, Synthetic aperture based on plenoptic camera for seeing through occlusions, in *Proceedings of the Advances in Multimedia Information Processing–PCM 2018* (Springer, 2018), pp. 158–167.
23. T. Yang, W. Ma, S. Wang, J. Li, J. Yu, Y. Zhang, Kinect based real-time synthetic aperture imaging through occlusion. *Multimed. Tools Appl.* **75**, 6925–6943 (2016).
24. N. Joshi, S. Avidan, W. Matusik, D. J. Kriegman, Synthetic aperture tracking: Tracking through occlusions, in *Proceedings of the 2007 IEEE 11th International Conference on Computer Vision*, Rio de Janeiro, Brazil, 14 to 21 October 2007, pp. 1–8.
25. Z. Pei, Y. Li, M. Ma, J. Li, C. Leng, X. Zhang, Y. Zhang, Occluded object 3D reconstruction using camera array synthetic aperture imaging. *Sensors* **19**, 607 (2019).
26. T. Yang, Y. Zhang, J. Yu, J. Li, W. Ma, X. Tong, R. Yu, L. Ran, All-in-focus synthetic aperture imaging, in *Proceedings of the Computer Vision–ECCV 2014* (Springer, 2014), pp. 1–15.
27. Z. Pei, Y. Zhang, X. Chen, Y.-H. Yang, Synthetic aperture imaging using pixel labeling via energy minimization. *Pattern Recogn.* **46**, 174–187 (2013).
28. I. Kurmi, D. C. Schedl, O. Bimber, Airborne optical sectioning. *J. Imag.* **4**, 102 (2018).
29. O. Bimber, I. Kurmi, D. C. Schedl, Synthetic aperture imaging with drones. *IEEE Comput. Graph. Appl.* **39**, 8–15 (2019).
30. I. Kurmi, D. C. Schedl, O. Bimber, A statistical view on synthetic aperture imaging for occlusion removal. *IEEE Sensors J.* **19**, 9374–9383 (2019).
31. I. Kurmi, D. C. Schedl, O. Bimber, Thermal airborne optical sectioning. *Remote Sens.* **11**, 1668 (2019).
32. D. C. Schedl, I. Kurmi, O. Bimber, Airborne optical sectioning for nesting observation. *Sci. Rep.* **10**, 7254 (2020).
33. I. Kurmi, D. C. Schedl, O. Bimber, Fast automatic visibility optimization for thermal synthetic aperture visualization. *IEEE Geosci. Remote Sens. Lett.* **18**, 836–840 (2021).
34. D. C. Schedl, I. Kurmi, O. Bimber, Search and rescue with airborne optical sectioning. *Nat. Mach. Intell.* **2**, 783–790 (2020).
35. G. Wetzstein, I. Ihrke, D. Lanman, W. Heidrich, Computational plenoptic imaging. *Comput. Graph. Forum.* **30**, 2397–2426 (2011).
36. G. Wu, B. Masia, A. Jarabo, Y. Zhang, L. Wang, Q. Dai, T. Chai, Y. Liu, Light field image processing: An overview. *IEEE J. Sel. Top. Signal Process.* **11**, 926–954 (2017).
37. E. H. Synge, A method of investigating the higher atmosphere. *London, Edinburgh, Dublin Philos. Mag. J. Sci.* **9**, 1014–1020 (1930).
38. V. Molebny, P. F. McManamon, O. Steinvall, T. Kobayashi, W. Chen, Laser radar: Historical prospective—From the east to the west. *Opt. Eng.* **56**, 031220 (2017).
39. B. Du, C. Pang, D. Wu, Z. Li, H. Peng, Y. Tao, E. Wu, G. Wu, High-speed photon-counting laser ranging for broad range of distances. *Sci. Rep.* **8**, 4198 (2018).
40. A. Bochkovskiy, C. Y. Wang, H. Y. M. Liao, YOLOv4: Optimal speed and accuracy of object detection (2020); https://arxiv.org/abs/2004.10934.
41. K. He, X. Zhang, S. Ren, J. Sun, Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **37**, 1904–1916 (2015).
42. Z. Huang, J. Wang, X. Fu, T. Yu, Y. Guo, R. Wang, DC-SPP-YOLO: Dense connection and spatial pyramid pooling based YOLO for object detection. *Inf. Sci.* **522**, 241–258 (2020).
43. R. Padilla, S. L. Netto, E. A. B. da Silva, A survey on performance metrics for object-detection algorithms, in *Proceedings of the 2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, Niterói, Brazil, 1 to 3 July 2020, pp. 237–242.
44. J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 27 to 30 June 2016, pp. 779–788.
45. V. S. Juan, M. Santos, J. M. Andújar, Intelligent UAV map generation and discrete path planning for search and rescue operations. *Complexity* **2018**, 6879419 (2018).
46. M. Radmanesh, M. Kumar, P. H. Guentert, M. Sarim, Overview of path-planning and obstacle avoidance algorithms for UAVs: A comparative study. *Unmanned Syst.* **6**, 95–118 (2018).
47. Z. Mengying, W. Hua, C. Feng, Online path planning algorithms for unmanned air vehicle, in *Proceedings of the 2017 IEEE International Conference on Unmanned Systems (ICUS)*, Beijing, China, 27 to 29 October 2017, pp. 116–119.
48. Y. Chen, G. Luo, Y. Mei, J. Yu, X. Su, UAV path planning using artificial potential field method updated by optimal control theory. *Int. J. Syst. Sci.* **47**, 1407–1420 (2016).
49. H. Khandani, H. Moradi, J. Y. Panah, A real-time coverage and tracking algorithm for UAVs based on potential field, in *Proceedings of the 2014 Second RSI/ISM International Conference on Robotics and Mechatronics (ICRoM)*, Tehran, Iran, 15 to 17 October 2014, pp. 700–705.
50. A. C. Woods, H. M. La, Dynamic target tracking and obstacle avoidance using a drone, in *Proceedings of the International Symposium on Visual Computing*, Las Vegas, NV, USA, 18 December 2015, pp. 857–866.
51. J. P. Queralta, J. Taipalmaa, B. C. Pullinen, V. K. Sarker, T. N. Gia, H. Tenhunen, M. Gabbouj, J. Raitoharju, T. Westerlund, Collaborative multi-robot search and rescue: Planning, coordination, perception, and active vision. *IEEE Access* **8**, 191617–191643 (2020).
52. S. Aggarwal, N. Kumar, Path planning techniques for unmanned aerial vehicles: A review, solutions, and challenges. *Comput. Commun.* **149**, 270–299 (2020).
53. T. M. Cabreira, L. B. Brisolara, P. R. Ferreira Jr., Survey on coverage path planning with unmanned aerial vehicles. *Drones* **3**, (2019).
54. G. Enric, M. Carreras, A survey on coverage path planning for robotics. *Robot. Auton. Syst.* **61**, 1258–1276 (2013).

Downloaded from https://www.science.org on October 19, 2022

55. F. C. J. Allaire, G. Labonté, M. Tarbouchi, V. Roberge, Recent advances in unmanned aerial vehicles real-time trajectory planning. *J. Unmanned Veh. Syst.* **7**, 259–295 (2019).
56. L. Yang, J. Qi, J. Xiao, X. Yong, A literature review of UAV 3D path planning, in *Proceedings of the 11th World Congress on Intelligent Control and Automation*, Shenyang, 29 June to 4 July 2014, pp. 2376–2381.
57. L. Quan, L. Han, B. Zhou, S. Shen, F. Gao, Survey of UAV motion planning. *IET Cyber Syst. Robot.* **2**, 14–21 (2020).
58. C. Goerzen, Z. Kong, B. Mettler, A survey of motion planning algorithms from the perspective of autonomous UAV guidance. *J. Intell. Robot. Syst.* **57**, 65–100 (2010).
59. I. Khoufi, A. Laouiti, C. Adjih, A survey of recent extended variants of the traveling salesman and vehicle routing problems for unmanned aerial vehicles. *Drones* **3**, 66 (2019).
60. W. P. Coutinho, M. Battarra, J. Fliege, The unmanned aerial vehicle routing and trajectory optimisation problem, a taxonomic review. *Comput. Ind. Eng.* **120**, 116–128 (2018).
61. A. Otto, N. Agatz, J. Campbell, B. Golden, E. Pesch, Optimization approaches for civil applications of unmanned aerial vehicles (UAVs) or aerial drones: A survey. *Networks* **72**, 411–458 (2018).
62. I. Kurmi, D. C. Schedl, O. Bimber, Pose error reduction for focus enhancement in thermal synthetic aperture visualization. *IEEE Geosci. Remote Sens. Lett.* 10.1109/LGRS.2021.3051718 (2021).
63. S. Hwang, J. Park, N. Kim, Y. Choi, I. S. Kweon, Multispectral pedestrian detection: Benchmark dataset and baselines, in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* Boston, MA, USA, 7 to 12 June 2015, pp. 1037–1045.
64. L. Lin, M. A. Goodrich, A Bayesian approach to modeling lost person behaviors based on terrain features in Wilderness Search and Rescue. *Comput. Math. Organ. Theory* **16**, 300–323 (2010).
65. A. Macwan, G. Nejat, B. Benhabib, Target-motion prediction for robotic search and rescue in wilderness environments. *IEEE Trans. Syst. Man Cybern. B Cybern.* **41**, 1287–1298 (2011).
66. R. D. Arnold, H. Yamaguchi, T. Tanaka, Search and rescue with autonomous flying robots through behavior-based cooperative intelligence. *Int. J. Humanitarian Action* **3**, 18 (2018).
67. E. Sava, C. Twardy, R. Koester, M. Sonwalkar, Evaluating lost person behavior models. *Trans. GIS* **20**, 38–53 (2016).
68. M. Segal, C. Korobkin, R. v. Widenfelt, J. Foran, P. Haeberli, Fast shadows and lighting effects using texture mapping. *SIGGRAPH Comput. Graph.* **26**, 249–252 (1992).
69. Upper Austrian government, digital elevation model of upper Austria; www.land-oberoesterreich.gv.at/211780.htm.
70. GDAL/OGR Geospatial Data Abstraction software Library. Open Source Geospatial Foundation (2021); https://gdal.org.
71. FLARM System Design and Interoperability (Technical Report, FLARM Technology Ltd, Switzerland, 2019); https://flarm.com/media/white-papers/interoperability/.
72. Avionics Supporting Automatic Dependent Surveillance–Broadcast (ADS-B) Aircraft Surveillance Applications (ASA) (TSO-C 195b, U.S. Department of Transportation, FAA, Washington, DC, 2014); www.faa.gov/nextgen/programs/adsb/.
73. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, L. Fei-Fei, ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis.* **115**, 211–252 (2015).
74. D. C. Schedl, I. Kurmi, O. Bimber, Data: Search and rescue with airborne optical sectioning, Version 2, Zenodo (2020); https://doi.org/10.5281/zenodo.3894773.
75. I. Kurmi, D. C. Schedl, O. Bimber, Data: Autonomous drones for search and rescue in forests, Version 1, Zenodo (2020); https://doi.org/10.5281/zenodo.4349220.

Funding: This research was funded by the Austrian Science Fund (FWF) under grant number P 32185-NBL and by the State of Upper Austria and the Austrian Federal Ministry of Education, Science, and Research via the LIT—Linz Institute of Technology under grant number LIT-2019-8-SEE-114. **Author contributions:** O.B. is the PI and originated the concept. O.B. and D.C.S. conceived and designed the experiments. D.C.S. and I.K. performed the experiments. O.B., D.C.S., and I.K. analyzed the data. D.C.S. and I.K. contributed materials/analysis tools. O.B., D.C.S., and I.K. wrote the paper. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The data collected in flights F12 to F28 can be downloaded from Zenodo (*74*) and includes GT labels, single images, integral images for evaluation, network configuration files, trained network weights, and classification results. Flights F0 to F11 are available at (*75*). We provide labels and integral images for the resampled 1D apertures (*74*). Scripts to compute Tables 1 and 2 are provided with the dataset. All source code and data are available at https://github.com/JKU-ICG/AOS. All other data needed to evaluate the conclusions of this paper can be found in the paper or the Supplementary Materials.

Submitted 21 December 2020
Accepted 28 May 2021
Published 23 June 2021
10.1126/scirobotics.abg1188

Citation: D. Schedl, I. Kurmi, O. Bimber, An autonomous drone for search and rescue in forests using airborne optical sectioning. *Sci. Robot.* **6**, eabg1188 (2021).

Downloaded from https://www.science.org on October 19, 2022

Science Robotics

An autonomous drone for search and rescue in forests using airborne optical sectioning

D. C. SchedlI. KurmiO. Bimber

Sci. Robot., 6 (55), eabg1188. • DOI: 10.1126/scirobotics.abg1188

View the article online
https://www.science.org/doi/10.1126/scirobotics.abg1188
Permissions
https://www.science.org/help/reprints-and-permissions

Downloaded from https://www.science.org on October 19, 2022

Use of this article is subject to the Terms of service

Science Robotics (ISSN 2470-9476) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Robotics* is a registered trademark of AAAS.
Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works